

RAW SEQUENCE LISTING  
PATENT APPLICATION US/08/848,439

TEAM-8

DATE: 06/30/97  
TIME: 14:29:20

INPUT SET: S18704.raw

This Raw Listing contains the General  
Information Section and up to the first 5 pages.

## SEQUENCE LISTING

ENTERED

## (1) General Information:

(i) APPLICANT: LaVALLIE, EDWARD  
RACIE, LISA

(ii) TITLE OF INVENTION: HUMAN SDF-5 PROTEIN AND COMPOSITIONS

(iii) NUMBER OF SEQUENCES: 3

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: GENETICS INSTITUTE, INC.  
(B) STREET: 87 CAMBRIDGE PARK DRIVE  
(C) CITY: CAMBRIDGE  
(D) STATE: MA  
(E) COUNTRY: USA  
(F) ZIP: 02140

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: LAZAR, STEVEN R.  
(B) REGISTRATION NUMBER: 32,618

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (617) 498-8260  
(B) TELEFAX: (617) 876-5851

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2027 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

RAW SEQUENCE LISTING  
PATENT APPLICATION US/08/848,439DATE: 06/30/97  
TIME: 14:29:22

INPUT SET: S18704.raw

47  
48 (ii) MOLECULE TYPE: DNA (genomic)  
49  
50  
51  
52 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:  
53  
54 GAATTCGGCC TTCATGGCCT AGCTCATTCT GCTCCCCCGG GTCGGAGCCC CCCGGAGCTG 60  
55  
56 CGCGCGGGCT TGCAGCGCCT CGCCCGCGCT CCTCCCGGTG TCCCGCTTCT CCGCGCCCCA 120  
57  
58 GCCGCCGGCT GCCAGCTTTT CGGGGCCCCG AGTCGCACCC AGCGAAGAGA GCGGGCCCCG 180  
59  
60 GACAAGCTCG AACTCCGGCC GCCTCGCCCT TCCCCGGCTC CGCTCCCTCT GCCCCCTCGG 240  
61  
62 GGTGCGCGCG CCACGATGCT GCAGGGCCCT GGCTCGCTGC TGCTGCTCTT CCTCGCCTCG 300  
63  
64 CACTGCTGCC TGGGCTCGGC GCGCGGGCTC TTCTCTTTG GCCAGCCCGA CTTCTCCTAC 360  
65  
66 AAGCGCAGCA ATTGCAAGCC CATCCCGGCC AACCTGCAGC TGTGCCACGG CATCGAATAC 420  
67  
68 CAGAACATGC GGCTGCCCAA CCTGCTGGGC CACGAGACCA TGAAGGAGGT GCTGGAGCAG 480  
69  
70 GCCGGCGCTT GGATCCCGCT GGTGATGAAG CAGTGCCACC CGGACACCAA GAAGTTCCCTG 540  
71  
72 TGCTCGCTCT TCGCCCCCGT CTGCCTCGAT GACCTAGACG AGACCATCCA GCCATGCCAC 600  
73  
74 TCGCTCTGCG TGCAGGTGAA GGACCGCTGC GCCCCGGTCA TGTCCGCCTT CGGCTTCCCC 660  
75  
76 TGGCCCGACA TGCTTGAGTG CGACCGTTTC CCCCAGGACA ACGACCTTTG CATCCCCCTC 720  
77  
78 GCTAGCAGCG ACCACCTCCT GCCAGCCACC GAGGAAGCTC CAAAGGTATG TGAAGCCTGC 780  
79  
80 AAAAATAAAA ATGATGATGA CAACGACATA ATGGAAACGC TTTGTAAAAA TGATTTTGCA 840  
81  
82 CTGAAAATAA AAGTGAAGGA GATAACCTAC ATCAACCGAG ATACCAAAAT CATCCTGGAG 900  
83  
84 ACCAAGAGCA AGACCATTTA CAAGCTGAAC GGTGTGTCCG AAAGGGACCT GAAGAAATCG 960  
85  
86 GTGCTGTGGC TCAAAGACAG CTTGCAGTGC ACCTGTGAGG AGATGAACGA CATCAACGCG 1020  
87  
88 CCCTATCTGG TCATGGGACA GAAACAGGGT GGGGAGCTGG TGATCACCTC GGTGAAGCGG 1080  
89  
90 TGGCAGAAGG GGCAGAGAGA GTTCAAGCGC ATCTCCCGCA GCATCCGCAA GCTGCAGTGC 1140  
91  
92 TAGTCCCGGC ATCCTGATGG CTCCGACAGG CCTGCTCCAG AGCACGGCTG ACCATTTCTG 1200  
93  
94 CTCCGGGATC TCAGCTCCCG TTCCCCAAGC AACTCCTAG CTGCTCCAGT CTCAGCCTGG 1260  
95  
96 GCAGCTTCCC CCTGCCTTTT GCACGTTTGC ATCCCCAGCA TTTCTGAGT TATAAGGCCA 1320  
97  
98 CAGGAGTGGA TAGCTGTTTT CACCTAAAGG AAAAGCCCAC CCGAATCTTG TAGAAATATT 1380  
99

# RAW SEQUENCE LISTING PATENT APPLICATION US/08/848,439

DATE: 06/30/97  
TIME: 14:29:24

INPUT SET: S18704.raw

```

100  CAAACTAATA AAATCATGAA TATTTTATG AAGTTTAAAA ATAGCTCACT TTAAAGCTAG 1440
101
102  TTTTGAATAG GTGCAACTGT GACTTGGGTC TGGTTGGTTG TTGTTTGTG TTTTGAGTCA 1500
103
104  GCTGATTTTC ACTTCCCACT GAGGTTGTCA TAACATGCAA ATTGCTTCAA TTTTCTCTGT 1560
105
106  GGCCCCAACT TGTGGGTCAC AAACCCTGTT GAGATAAAGC TGGCTGTTAT CTCAACATCT 1620
107
108  TCATCAGCTC CAGACTGAGA CTCAGTGTCT AAGTCTTACA ACAATTCATC ATTTTATACC 1680
109
110  TTCAATGGGA ACTTAAACTG TTACATGTAT CACATTCCAG CTACAATACT TCCATTTATT 1740
111
112  AGAAGCACAT TAACCATTTC TATAGCATGA TTTCTTCAAG TAAAAGGCAA AAGATATAAA 1800
113
114  TTTTATAATT GACTTGAGTA CTTTAAGCCT TGTTTAAAAC ATTTCTTACT TAACTTTTGC 1860
115
116  AAATTAAACC CATTGTAGCT TACCTGTAAT ATACATAGTA GTTTACCTTT AAAAGTTGTA 1920
117
118  AAAATATTGC TTTAACCAAC ACTGTAAATA TTTCAGATAA ACATTATATT CTTGTATATA 1980
119
120  AACTTTACAT CCTGTTTAC CTAAAAAAG AAAAAAAG CGGCCGC 2027
121

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 295 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

136  Met Leu Gln Gly Pro Gly Ser Leu Leu Leu Leu Phe Leu Ala Ser His
137  1             5             10             15
138
139  Cys Cys Leu Gly Ser Ala Arg Gly Leu Phe Leu Phe Gly Gln Pro Asp
140  20             25             30
141
142  Phe Ser Tyr Lys Arg Ser Asn Cys Lys Pro Ile Pro Ala Asn Leu Gln
143  35             40             45
144
145  Leu Cys His Gly Ile Glu Tyr Gln Asn Met Arg Leu Pro Asn Leu Leu
146  50             55             60
147
148  Gly His Glu Thr Met Lys Glu Val Leu Glu Gln Ala Gly Ala Trp Ile
149  65             70             75             80
150
151  Pro Leu Val Met Lys Gln Cys His Pro Asp Thr Lys Lys Phe Leu Cys
152  85             90             95

```

RAW SEQUENCE LISTING  
PATENT APPLICATION US/08/848,439DATE: 06/30/97  
TIME: 14:29:26

INPUT SET: S18704.raw

153  
154 Ser Leu Phe Ala Pro Val Cys Leu Asp Asp Leu Asp Glu Thr Ile Gln  
155 100 105 110  
156  
157 Pro Cys His Ser Leu Cys Val Gln Val Lys Asp Arg Cys Ala Pro Val  
158 115 120 125  
159  
160 Met Ser Ala Phe Gly Phe Pro Trp Pro Asp Met Leu Glu Cys Asp Arg  
161 130 135 140  
162  
163 Phe Pro Gln Asp Asn Asp Leu Cys Ile Pro Leu Ala Ser Ser Asp His  
164 145 150 155 160  
165  
166 Leu Leu Pro Ala Thr Glu Glu Ala Pro Lys Val Cys Glu Ala Cys Lys  
167 165 170 175  
168  
169 Asn Lys Asn Asp Asp Asp Asn Asp Ile Met Glu Thr Leu Cys Lys Asn  
170 180 185 190  
171  
172 Asp Phe Ala Leu Lys Ile Lys Val Lys Glu Ile Thr Tyr Ile Asn Arg  
173 195 200 205  
174  
175 Asp Thr Lys Ile Ile Leu Glu Thr Lys Ser Lys Thr Ile Tyr Lys Leu  
176 210 215 220  
177  
178 Asn Gly Val Ser Glu Arg Asp Leu Lys Lys Ser Val Leu Trp Leu Lys  
179 225 230 235 240  
180  
181 Asp Ser Leu Gln Cys Thr Cys Glu Glu Met Asn Asp Ile Asn Ala Pro  
182 245 250 255  
183  
184 Tyr Leu Val Met Gly Gln Lys Gln Gly Gly Glu Leu Val Ile Thr Ser  
185 260 265 270  
186  
187 Val Lys Arg Trp Gln Lys Gly Gln Arg Glu Phe Lys Arg Ile Ser Arg  
188 275 280 285  
189  
190 Ser Ile Arg Lys Leu Gln Cys  
191 290 295  
192

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 275 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

**INPUT SET: S18704.raw**

206																	
207	Ser	Ala	Arg	Gly	Leu	Phe	Leu	Phe	Gly	Gln	Pro	Asp	Phe	Ser	Tyr	Lys	
208	1				5					10					15		
209																	
210	Arg	Ser	Asn	Cys	Lys	Pro	Ile	Pro	Ala	Asn	Leu	Gln	Leu	Cys	His	Gly	
211				20					25					30			
212																	
213	Ile	Glu	Tyr	Gln	Asn	Met	Arg	Leu	Pro	Asn	Leu	Leu	Gly	His	Glu	Thr	
214			35					40					45				
215																	
216	Met	Lys	Glu	Val	Leu	Glu	Gln	Ala	Gly	Ala	Trp	Ile	Pro	Leu	Val	Met	
217		50					55					60					
218																	
219	Lys	Gln	Cys	His	Pro	Asp	Thr	Lys	Lys	Phe	Leu	Cys	Ser	Leu	Phe	Ala	
220	65					70					75					80	
221																	
222	Pro	Val	Cys	Leu	Asp	Asp	Leu	Asp	Glu	Thr	Ile	Gln	Pro	Cys	His	Ser	
223					85					90					95		
224																	
225	Leu	Cys	Val	Gln	Val	Lys	Asp	Arg	Cys	Ala	Pro	Val	Met	Ser	Ala	Phe	
226				100					105					110			
227																	
228	Gly	Phe	Pro	Trp	Pro	Asp	Met	Leu	Glu	Cys	Asp	Arg	Phe	Pro	Gln	Asp	
229			115					120					125				
230																	
231	Asn	Asp	Leu	Cys	Ile	Pro	Leu	Ala	Ser	Ser	Asp	His	Leu	Leu	Pro	Ala	
232		130					135					140					
233																	
234	Thr	Glu	Glu	Ala	Pro	Lys	Val	Cys	Glu	Ala	Cys	Lys	Asn	Lys	Asn	Asp	
235	145					150					155					160	
236																	
237	Asp	Asp	Asn	Asp	Ile	Met	Glu	Thr	Leu	Cys	Lys	Asn	Asp	Phe	Ala	Leu	
238					165					170					175		
239																	
240	Lys	Ile	Lys	Val	Lys	Glu	Ile	Thr	Tyr	Ile	Asn	Arg	Asp	Thr	Lys	Ile	
241				180					185					190			
242																	
243	Ile	Leu	Glu	Thr	Lys	Ser	Lys	Thr	Ile	Tyr	Lys	Leu	Asn	Gly	Val	Ser	
244			195					200					205				
245																	
246	Glu	Arg	Asp	Leu	Lys	Lys	Ser	Val	Leu	Trp	Leu	Lys	Asp	Ser	Leu	Gln	
247		210					215					220					
248																	

PAGE: 1

**SEQUENCE VERIFICATION REPORT**  
**PATENT APPLICATION US/08/848,439**

DATE: 06/30/97  
TIME: 14:29:32

*INPUT SET: S18704.raw*

Line

Error

Original Text